Human Pose Estimation

Minju Kim

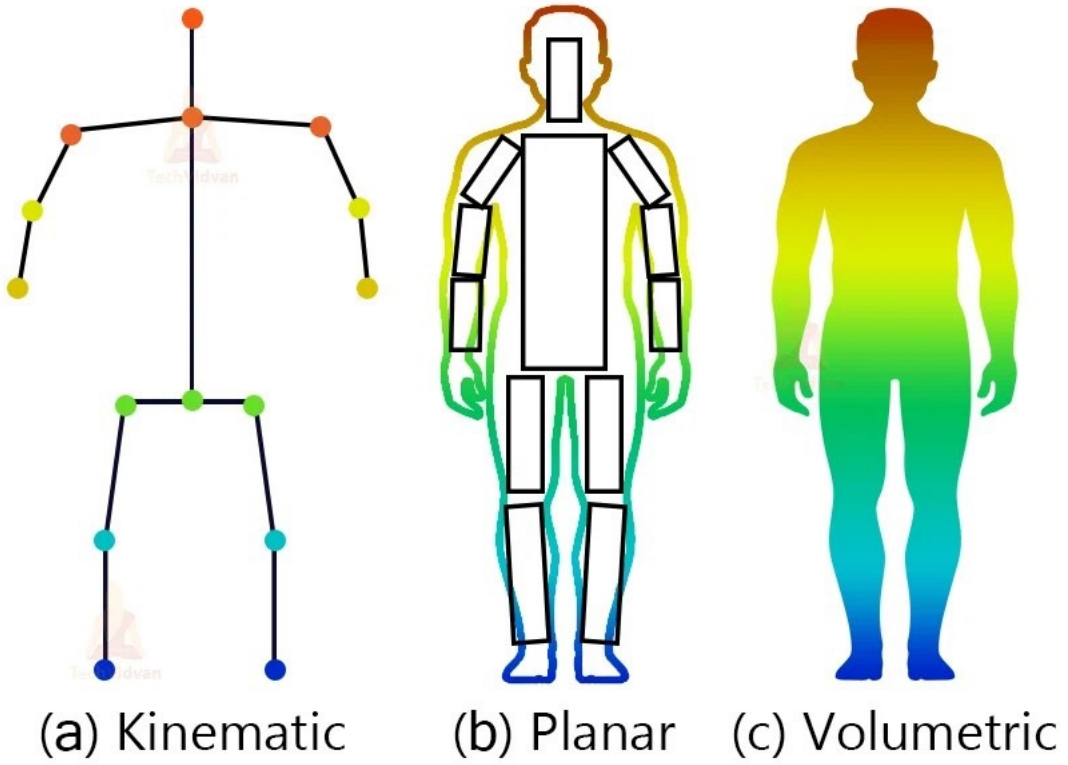
2022.03.26

What is Human Pose Estimation?

- Human Pose estimation is a computer vision task that represents the orientation of a person in a graphical format.
- Human Pose Estimation identifies and classifies the poses of human body parts and joints in images or videos.
- Essentially it is a way to capture a set of coordinates by defining the human body joints like wrist, shoulder, knees, eyes, ears, ankles, and arms, which is a key point in images and videos that can describe a pose of a person.
- When an image or video is given to the pose estimator model as input, it identifies the coordinates of these detected body parts and joints as output and a confidence score showing precision of the estimations.

Types of Human Pose Estimation Models

- (a) Skeleton-based model :
  - also called the kinematic model
  - this representative includes a set of key points (joints) like ankles, knees, shoulders, elbows, wrists, and limb orientations primarily utilized for 3D and 2D pose estimation.
- (b) Countour-based model :
  - also called the planar model, used for 2D pose estimation
  - It represents the appearance and shape of a human body, where body parts are displayed with boundaries and rectangles of a person's contour.
- (c) Volume-based model :
  - also called the volumetric model, is used for 3D pose estimation
  - it consists of multiple popular 3D human body models and poses represented by human geometric meshes and shape



Human Pose Estimation using Deep Neural Networks

Top-down

- It run a body detector first and determine body joints within the discovered bounding boxes.
- The top-down approach to HPE raises a lot of error in localization and inaccuracies during prediction and is, therefore, quite challenging.

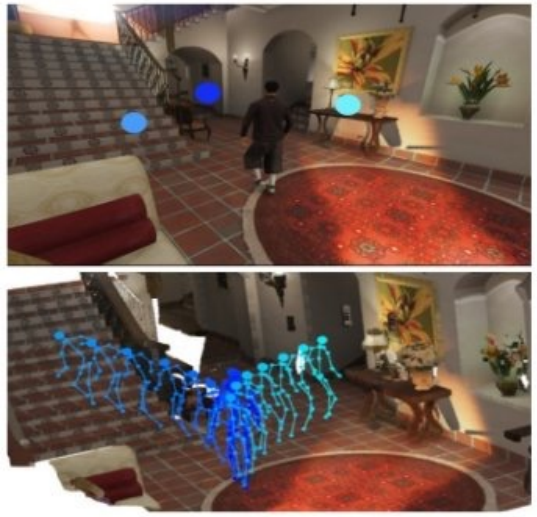
Bottom-up

- It evaluates each body joint first and then arrange them to compose a unique pose.
- Estimate the human body parts in the image followed by calculating the pose.

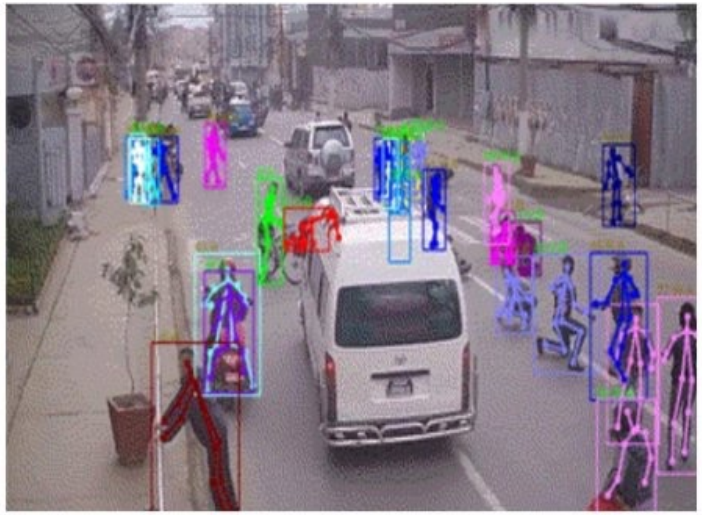
AlphaPose (RMPE, **Regional Multi-person Pose Estimation**)

- It is a popular top-down method of pose estimation.
- It is an optimal architecture for estimating human poses via optimally detected bounding boxes.
- AlphaPose architecture is applicable for detecting both single and multi-person poses in images or video fields.

Applications of HPE



Action prediction

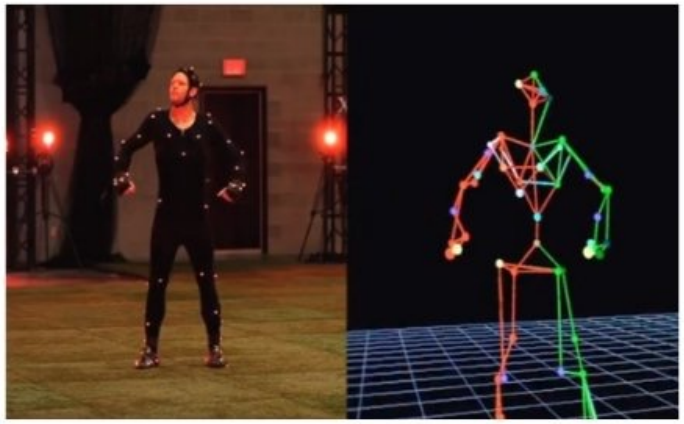


Surveillance

Cloth Parsing

Online Coaching

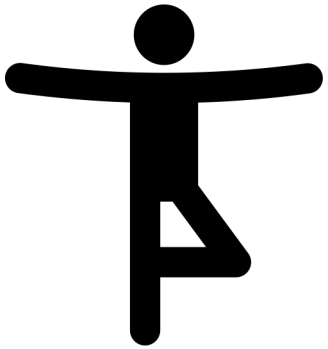
Movie and Game



AR and VR

Healthcare

Applications of HPE

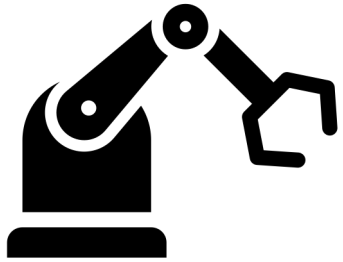


1. Human Activity & Movement Estimation

- AI-powered sports coaches or personal gym trainer
- Sign language communication for disabled
- Monitoring movements in security and surveillance

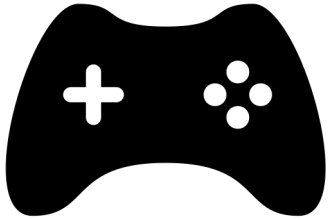
2. Augmented Reality & Virtual Reality (AR/VR)

- When clubbed with augmented and virtual reality applications, human pose estimation presents an opportunity to create more realistic and responsive experiences.



3. Robotics

- More responsive, flexible, and true-to-life robotics systems.



4. Animation & Gaming

- Modern advancements in pose estimation and motion capture technology make character animation a streamlined and automated process.

Referneces

- https://www.v7labs.com/blog/human-pose-estimation-guide
- https://www.analyticsvidhya.com/blog/2022/01/a-comprehensive-guide-on-human-pose-estimation/
- https://www.youtube.com/watch?v=_sobpAW16c0
- https://viso.ai/deep-learning/pose-estimation-ultimate-overview/
- https://github.com/MVIG-SJTU/AlphaPose/tree/master
- https://nanonets.com/blog/human-pose-estimation-2d-guide/

Thank you